In [2]:

```
!pip install piml
```

```
Requirement already satisfied: piml in /Users/jakebrulato/Documents/GitHub/DSBA6010/.venv
/lib/python3.10/site-packages (0.6.0.post2)
Requirement already satisfied: ipykernel in /Users/jakebrulato/Documents/GitHub/DSBA6010/
.venv/lib/python3.10/site-packages (from piml) (6.29.5)
Requirement already satisfied: ipywidgets>=7.7.0 in /Users/jakebrulato/Documents/GitHub/D
SBA6010/.venv/lib/python3.10/site-packages (from piml) (8.1.5)
Requirement already satisfied: joblib>=1.1.0 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from piml) (1.4.2)
Requirement already satisfied: ipython>=7.12.0 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from piml) (8.27.0)
Requirement already satisfied: numpy<1.24.0,>=1.21.4 in /Users/jakebrulato/Documents/GitH
ub/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (1.23.5)
Requirement already satisfied: scipy<=1.10.1,>=1.5.3 in /Users/jakebrulato/Documents/GitH
ub/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (1.10.1)
Requirement already satisfied: pandas<2.0.0,>=1.3.5 in /Users/jakebrulato/Documents/GitHu
b/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (1.5.3)
Requirement already satisfied: matplotlib<3.8.0,>=3.1.2 in /Users/jakebrulato/Documents/G
itHub/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (3.7.5)
Requirement already satisfied: seaborn>=0.11.2 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from piml) (0.13.2)
Requirement already satisfied: xlrd>=1.2.0 in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from piml) (2.0.1)
Requirement already satisfied: scikit-learn<1.4.0,>=0.24.2 in /Users/jakebrulato/Document
s/GitHub/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (1.3.2)
Requirement already satisfied: xgboost>=1.4.2 in /Users/jakebrulato/Documents/GitHub/DSBA
6010/.venv/lib/python3.10/site-packages (from piml) (2.1.1)
Requirement already satisfied: statsmodels>=0.12.2 in /Users/jakebrulato/Documents/GitHub
/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (0.14.2)
Requirement already satisfied: lime>=0.2.0.1 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from piml) (0.2.0.1)
Requirement already satisfied: shap>=0.39.0 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from piml) (0.46.0)
Requirement already satisfied: torch>=1.11.0 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from piml) (2.2.2)
Requirement already satisfied: pygam==0.8.0 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from piml) (0.8.0)
Requirement already satisfied: natsort>=8.2.0 in /Users/jakebrulato/Documents/GitHub/DSBA
6010/.venv/lib/python3.10/site-packages (from piml) (8.4.0)
Requirement already satisfied: psutil>=5.9.0 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from piml) (6.0.0)
Requirement already satisfied: dill>=0.3.6 in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from piml) (0.3.8)
Requirement already satisfied: packaging>=20.5 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from piml) (24.1)
Requirement already satisfied: networkx>=2.6.3 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from piml) (3.3)
Requirement already satisfied: numba<0.57.0 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from piml) (0.56.4)
Requirement already satisfied: jupyter-client<=7.4.9 in /Users/jakebrulato/Documents/GitH
ub/DSBA6010/.venv/lib/python3.10/site-packages (from piml) (7.4.9)
Requirement already satisfied: optbinning>=0.17.3 in /Users/jakebrulato/Documents/GitHub/
DSBA6010/.venv/lib/python3.10/site-packages (from piml) (0.19.0)
Requirement already satisfied: momentchi2 in /Users/jakebrulato/Documents/GitHub/DSBA6010
/.venv/lib/python3.10/site-packages (from piml) (0.1.8)
Requirement already satisfied: future in /Users/jakebrulato/Documents/GitHub/DSBA6010/.ve
nv/lib/python3.10/site-packages (from pygam==0.8.0->piml) (1.0.0)
Requirement already satisfied: progressbar2 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from pygam==0.8.0->piml) (4.5.0)
Requirement already satisfied: decorator in /Users/jakebrulato/Documents/GitHub/DSBA6010/
.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (5.1.1)
Requirement already satisfied: jedi>=0.16 in /Users/jakebrulato/Documents/GitHub/DSBA6010
/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (0.19.1)
Requirement already satisfied: matplotlib-inline in /Users/jakebrulato/Documents/GitHub/D
SBA6010/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (0.1.7)
Requirement already satisfied: prompt-toolkit<3.1.0,>=3.0.41 in /Users/jakebrulato/Docume
```

```
nts/GitHub/DSBA6010/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (3.0.
47)
Requirement already satisfied: pygments>=2.4.0 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (2.18.0)
Requirement already satisfied: stack-data in /Users/jakebrulato/Documents/GitHub/DSBA6010
/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (0.6.3)
Requirement already satisfied: traitlets>=5.13.0 in /Users/jakebrulato/Documents/GitHub/D
SBA6010/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (5.14.3)
Requirement already satisfied: exceptiongroup in /Users/jakebrulato/Documents/GitHub/DSBA
6010/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (1.2.2)
Requirement already satisfied: typing-extensions>=4.6 in /Users/jakebrulato/Documents/Git
Hub/DSBA6010/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (4.12.2)
Requirement already satisfied: pexpect>4.3 in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from ipython>=7.12.0->piml) (4.9.0)
Requirement already satisfied: comm>=0.1.3 in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from ipywidgets>=7.7.0->piml) (0.2.2)
Requirement already satisfied: widgetsnbextension~=4.0.12 in /Users/jakebrulato/Documents
/GitHub/DSBA6010/.venv/lib/python3.10/site-packages (from ipywidgets>=7.7.0->piml) (4.0.1
3)
Requirement already satisfied: jupyterlab-widgets~=3.0.12 in /Users/jakebrulato/Documents
/GitHub/DSBA6010/.venv/lib/python3.10/site-packages (from ipywidgets>=7.7.0->piml) (3.0.1
3)
Requirement already satisfied: entrypoints in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from jupyter-client<=7.4.9->piml) (0.4)
Requirement already satisfied: jupyter-core>=4.9.2 in /Users/jakebrulato/Documents/GitHub
/DSBA6010/.venv/lib/python3.10/site-packages (from jupyter-client<=7.4.9->piml) (5.7.2)
Requirement already satisfied: nest-asyncio>=1.5.4 in /Users/jakebrulato/Documents/GitHub
/DSBA6010/.venv/lib/python3.10/site-packages (from jupyter-client<=7.4.9->piml) (1.6.0)
Requirement already satisfied: python-dateutil>=2.8.2 in /Users/jakebrulato/Documents/Git
Hub/DSBA6010/.venv/lib/python3.10/site-packages (from jupyter-client<=7.4.9->piml) (2.9.0
.post0)
Requirement already satisfied: pyzmq>=23.0 in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from jupyter-client<=7.4.9->piml) (24.0.1)
Requirement already satisfied: tornado>=6.2 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from jupyter-client<=7.4.9->piml) (6.4.1)
Requirement already satisfied: tqdm in /Users/jakebrulato/Documents/GitHub/DSBA6010/.venv
/lib/python3.10/site-packages (from lime>=0.2.0.1->piml) (4.66.5)
Requirement already satisfied: scikit-image>=0.12 in /Users/jakebrulato/Documents/GitHub/
DSBA6010/.venv/lib/python3.10/site-packages (from lime>=0.2.0.1->piml) (0.24.0)
Requirement already satisfied: contourpy>=1.0.1 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from matplotlib<3.8.0,>=3.1.2->piml) (1.3.0)
Requirement already satisfied: cycler>=0.10 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from matplotlib<3.8.0,>=3.1.2->piml) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /Users/jakebrulato/Documents/GitHub/D
SBA6010/.venv/lib/python3.10/site-packages (from matplotlib<3.8.0,>=3.1.2->piml) (4.53.1)
Requirement already satisfied: kiwisolver>=1.0.1 in /Users/jakebrulato/Documents/GitHub/D
SBA6010/.venv/lib/python3.10/site-packages (from matplotlib<3.8.0,>=3.1.2->piml) (1.4.5)
Requirement already satisfied: pillow>=6.2.0 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from matplotlib<3.8.0,>=3.1.2->piml) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from matplotlib<3.8.0,>=3.1.2->piml) (3.1.4)
Requirement already satisfied: llvmlite<0.40,>=0.39.0dev0 in /Users/jakebrulato/Documents
/GitHub/DSBA6010/.venv/lib/python3.10/site-packages (from numba<0.57.0->piml) (0.39.1)
Requirement already satisfied: setuptools in /Users/jakebrulato/Documents/GitHub/DSBA6010
/.venv/lib/python3.10/site-packages (from numba<0.57.0->piml) (65.5.0)
Requirement already satisfied: ortools>=9.4 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from optbinning>=0.17.3->piml) (9.7.2996)
Requirement already satisfied: ropwr>=1.0.0 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from optbinning>=0.17.3->piml) (1.0.0)
Requirement already satisfied: pytz>=2020.1 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from pandas<2.0.0,>=1.3.5->piml) (2024.1)
Requirement already satisfied: threadpoolctl>=2.0.0 in /Users/jakebrulato/Documents/GitHu
b/DSBA6010/.venv/lib/python3.10/site-packages (from scikit-learn<1.4.0,>=0.24.2->piml) (3
.5.0)
Requirement already satisfied: slicer==0.0.8 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from shap>=0.39.0->piml) (0.0.8)
Requirement already satisfied: cloudpickle in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from shap>=0.39.0->piml) (3.0.0)
Requirement already satisfied: patsy>=0.5.6 in /Users/jakebrulato/Documents/GitHub/DSBA60
10/.venv/lib/python3.10/site-packages (from statsmodels>=0.12.2->piml) (0.5.6)
Requirement already satisfied: filelock in /Users/jakebrulato/Documents/GitHub/DSBA6010/.
venv/lib/python3.10/site-packages (from torch>=1.11.0->piml) (3.15.4)
```

```
Requirement already satisfied: sympy in /Users/jakebrulato/Documents/GitHub/DSBA6010/.ven
v/lib/python3.10/site-packages (from torch>=1.11.0->piml) (1.13.2)
Requirement already satisfied: jinja2 in /Users/jakebrulato/Documents/GitHub/DSBA6010/.ve
nv/lib/python3.10/site-packages (from torch>=1.11.0->piml) (3.1.4)
Requirement already satisfied: fsspec in /Users/jakebrulato/Documents/GitHub/DSBA6010/.ve
nv/lib/python3.10/site-packages (from torch>=1.11.0->piml) (2024.6.1)
Requirement already satisfied: appnope in /Users/jakebrulato/Documents/GitHub/DSBA6010/.v
env/lib/python3.10/site-packages (from ipykernel->piml) (0.1.4)
Requirement already satisfied: debugpy>=1.6.5 in /Users/jakebrulato/Documents/GitHub/DSBA
6010/.venv/lib/python3.10/site-packages (from ipykernel->piml) (1.8.5)
Requirement already satisfied: parso<0.9.0,>=0.8.3 in /Users/jakebrulato/Documents/GitHub
/DSBA6010/.venv/lib/python3.10/site-packages (from jedi>=0.16->ipython>=7.12.0->piml) (0.
8.4)
Requirement already satisfied: platformdirs>=2.5 in /Users/jakebrulato/Documents/GitHub/D
SBA6010/.venv/lib/python3.10/site-packages (from jupyter-core>=4.9.2->jupyter-client<=7.4
.9->piml) (4.2.2)
Requirement already satisfied: absl-py>=0.13 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from ortools>=9.4->optbinning>=0.17.3->piml) (2.1
.0)
Requirement already satisfied: protobuf>=4.23.3 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from ortools>=9.4->optbinning>=0.17.3->piml) (
4.25.5)
Requirement already satisfied: six in /Users/jakebrulato/Documents/GitHub/DSBA6010/.venv/
lib/python3.10/site-packages (from patsy>=0.5.6->statsmodels>=0.12.2->piml) (1.16.0)
Requirement already satisfied: ptyprocess>=0.5 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from pexpect>4.3->ipython>=7.12.0->piml) (0.7.0
)
Requirement already satisfied: wcwidth in /Users/jakebrulato/Documents/GitHub/DSBA6010/.v
env/lib/python3.10/site-packages (from prompt-toolkit<3.1.0,>=3.0.41->ipython>=7.12.0->pi
ml) (0.2.13)
Requirement already satisfied: cvxpy>=1.1.14 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from ropwr>=1.0.0->optbinning>=0.17.3->piml) (1.5
.3)
Requirement already satisfied: imageio>=2.33 in /Users/jakebrulato/Documents/GitHub/DSBA6
010/.venv/lib/python3.10/site-packages (from scikit-image>=0.12->lime>=0.2.0.1->piml) (2.
35.1)
Requirement already satisfied: tifffile>=2022.8.12 in /Users/jakebrulato/Documents/GitHub
/DSBA6010/.venv/lib/python3.10/site-packages (from scikit-image>=0.12->lime>=0.2.0.1->pim
l) (2024.8.28)
Requirement already satisfied: lazy-loader>=0.4 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from scikit-image>=0.12->lime>=0.2.0.1->piml)
(0.4)
Requirement already satisfied: MarkupSafe>=2.0 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from jinja2->torch>=1.11.0->piml) (2.1.5)
Requirement already satisfied: python-utils>=3.8.1 in /Users/jakebrulato/Documents/GitHub
/DSBA6010/.venv/lib/python3.10/site-packages (from progressbar2->pygam==0.8.0->piml) (3.8
.2)
Requirement already satisfied: executing>=1.2.0 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from stack-data->ipython>=7.12.0->piml) (2.0.1
)
Requirement already satisfied: asttokens>=2.1.0 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from stack-data->ipython>=7.12.0->piml) (2.4.1
)
Requirement already satisfied: pure-eval in /Users/jakebrulato/Documents/GitHub/DSBA6010/
.venv/lib/python3.10/site-packages (from stack-data->ipython>=7.12.0->piml) (0.2.3)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /Users/jakebrulato/Documents/GitHub/
DSBA6010/.venv/lib/python3.10/site-packages (from sympy->torch>=1.11.0->piml) (1.3.0)
Requirement already satisfied: osqp>=0.6.2 in /Users/jakebrulato/Documents/GitHub/DSBA601
0/.venv/lib/python3.10/site-packages (from cvxpy>=1.1.14->ropwr>=1.0.0->optbinning>=0.17.
3->piml) (0.6.7.post1)
Requirement already satisfied: ecos>=2 in /Users/jakebrulato/Documents/GitHub/DSBA6010/.v
env/lib/python3.10/site-packages (from cvxpy>=1.1.14->ropwr>=1.0.0->optbinning>=0.17.3->p
iml) (2.0.14)
Requirement already satisfied: clarabel>=0.5.0 in /Users/jakebrulato/Documents/GitHub/DSB
A6010/.venv/lib/python3.10/site-packages (from cvxpy>=1.1.14->ropwr>=1.0.0->optbinning>=0
.17.3->piml) (0.9.0)
Requirement already satisfied: scs>=3.2.4.post1 in /Users/jakebrulato/Documents/GitHub/DS
BA6010/.venv/lib/python3.10/site-packages (from cvxpy>=1.1.14->ropwr>=1.0.0->optbinning>=
0.17.3->piml) (3.2.7)
Requirement already satisfied: qdldl in /Users/jakebrulato/Documents/GitHub/DSBA6010/.ven
v/lib/python3.10/site-packages (from osqp>=0.6.2->cvxpy>=1.1.14->ropwr>=1.0.0->optbinning
>=0.17.3->piml) (0.1.7.post4)
```

Data Pipeline

In [3]:

```
from piml import Experiment
exp_tw_credit = Experiment()
exp_bike_sharing = Experiment()
exp_tw_credit.data_loader("TaiwanCredit", silent=True)
exp_bike_sharing.data_loader("BikeSharing", silent=True)
```

In [4]:

```
# Display data summary tables
tw_credit_summary = exp_tw_credit.data_summary()
bike_sharing_summary = exp_bike_sharing.data_summary()
```

Q1 (5pts): Using PiML, show data summary tables for both data sets. What is the max value of `cnt` in the bike share data? How many unique values are there in the `EDUCATION` data point of the Taiwan credit data?

- **#### There are 4 unique values for Education and 977 is the max value for 'cnt'**

Model Pipeline

In [5]:

```
from xgboost import XGBClassifier, XGBRegressor
from sklearn.neural_network import MLPClassifier, MLPRegressor

exp_bike_sharing.data_prepare(random_state=100, target="cnt", test_ratio=0.2, task_type=
"regression", split_method="random", silent=True)
exp_tw_credit.data_prepare(random_state=100, target="FlagDefault", test_ratio=0.2, task_t
ype="classification", split_method="random", silent=True)

exp_bike_sharing.model_train(model=XGBRegressor(max_depth=5, n_estimators=500), name="XG
B_Bike_Sharing")
exp_bike_sharing.model_train(model=MLPRegressor(hidden_layer_sizes=[10]*2, activation="r
elu", random_state=0, early_stopping=True), name="DNN_Bike_Sharing")

exp_tw_credit.model_train(model=XGBClassifier(max_depth=5, n_estimators=500), name="XGB_T
W_Credit")
exp_tw_credit.model_train(model=MLPClassifier(hidden_layer_sizes=[10]*2, activation="relu
", random_state=0, early_stopping=True), name="DNN_TW_Credit")
```

Q2 (5pts): What does the `random_state` parameter do? Will the XGB model for the bike sharing data turn out identical every time you run the code above?

- **#### The random_state parameter is used to create reproducibility in models and processes that involve randomness. By setting a fixed value for random_state, you make sure that the random processes involved (such as random splits in data or initialization of model weights) are performed in the same way every time the code is run. XGBoost has multiple areas that have randomness and random state cannot control those creating similar outputs.**

Q3 (2.5pts): What are the hyperparameters specified in each type of model?

- **#### XGB model hyperparameters are max_depth and n_estimators, MLP model hyperparameters are ' hidden_layer_sizes, activation function, random state, and early stopping,**

Bike sharing dataset with XGB model and DNN model

Q4 (10 pts): Compute the reliability table (including empirical coverage and average bandwidth) of the two Bike Sharing models; which model is more reliable?

In [6]:

```
# Evaluate the reliability for the XGBoost model
reliability_xgb = exp_bike_sharing.model_diagnose(model="XGB_Bike_Sharing", show="reliabi
lity_table")

# Evaluate the reliability for the DNN model
reliability_dnn = exp_bike_sharing.model_diagnose(model="DNN_Bike_Sharing", show="reliabi
lity_table")
```

	Empirical Coverage	Average Bandwidth
0	0.8813	0.1042

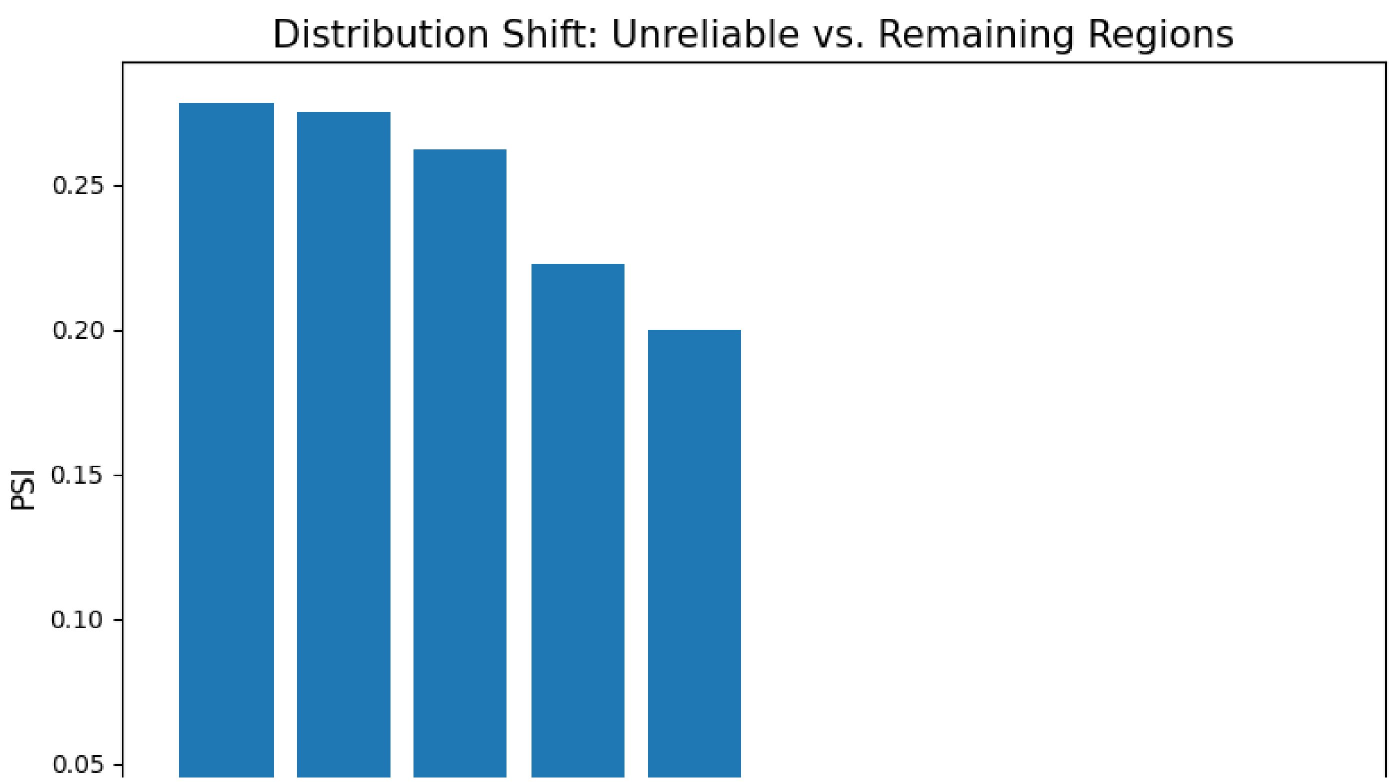
	Empirical Coverage	Average Bandwidth
0	0.8856	0.1874

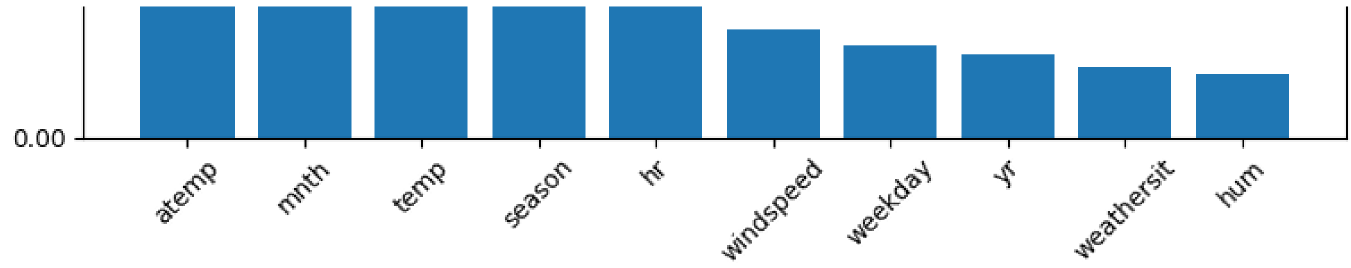
- **#### XGBoost has a more narrow bandwith, making it more likely to be reliable than the DNN.**

Q5 (5 pts): Use the `reliability_distance` function in PiML on the XGB model of the bike sharing data. List the top 5 features that have the largest distributional distance between unreliable regions and reliable regions based on PSI score. (You can use the defaults for all other arguments of `model_diagnose`)

In [7]:

```
exp_bike_sharing.model_diagnose(model="XGB_Bike_Sharing", show="reliability_distance")
```



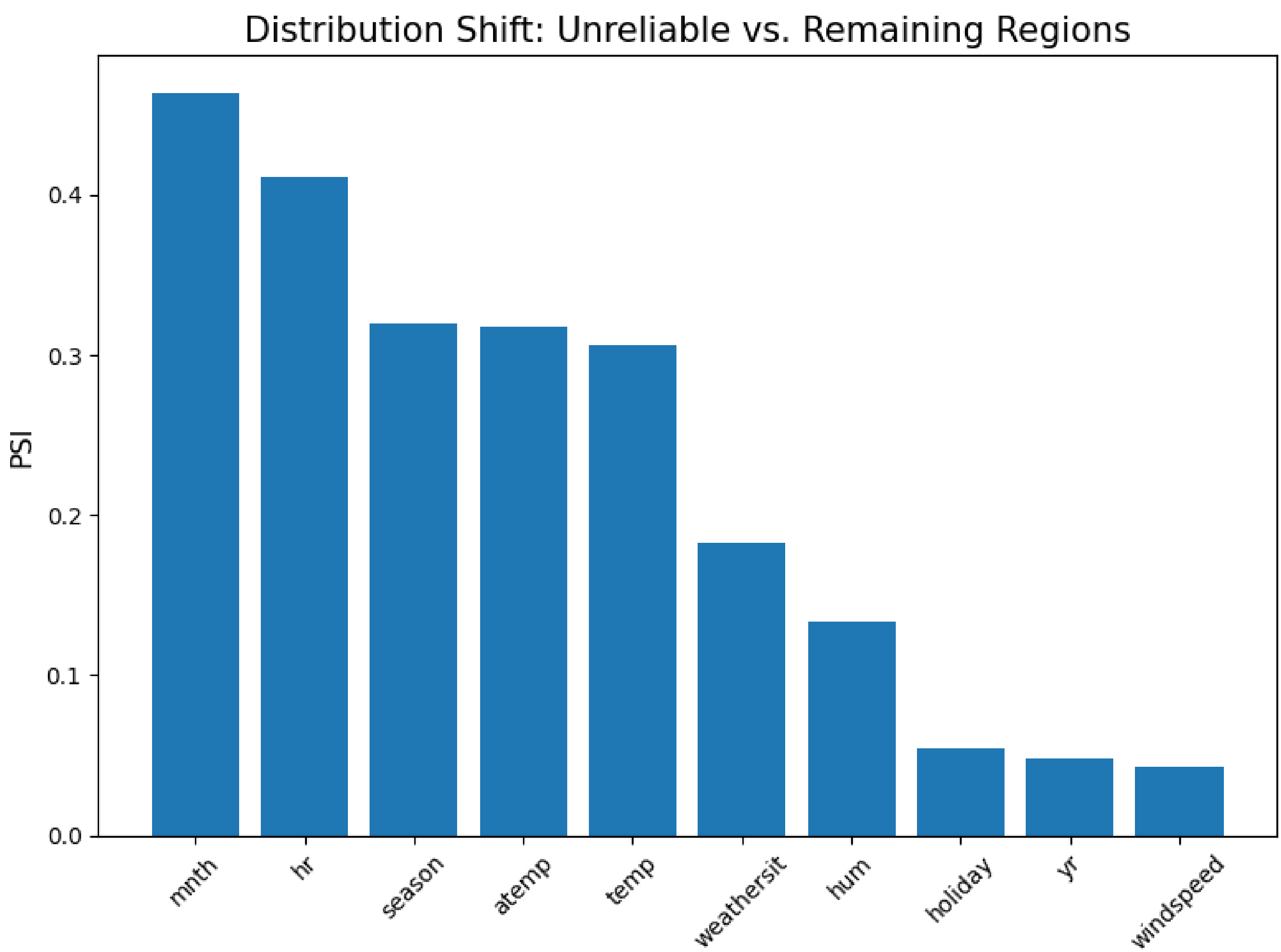


- ### Top 5 Features:
- #### XGB: atemp, mnth, temp, season, hr

Q6 (5 pts): Do the same analysis as in Q5, but use a threshold of 1.3. Why did the PSI values increase?

In [8]:

```
exp_bike_sharing.model_diagnose(model="XGB_Bike_Sharing", show="reliability_distance", th
reshold= 1.3)
```

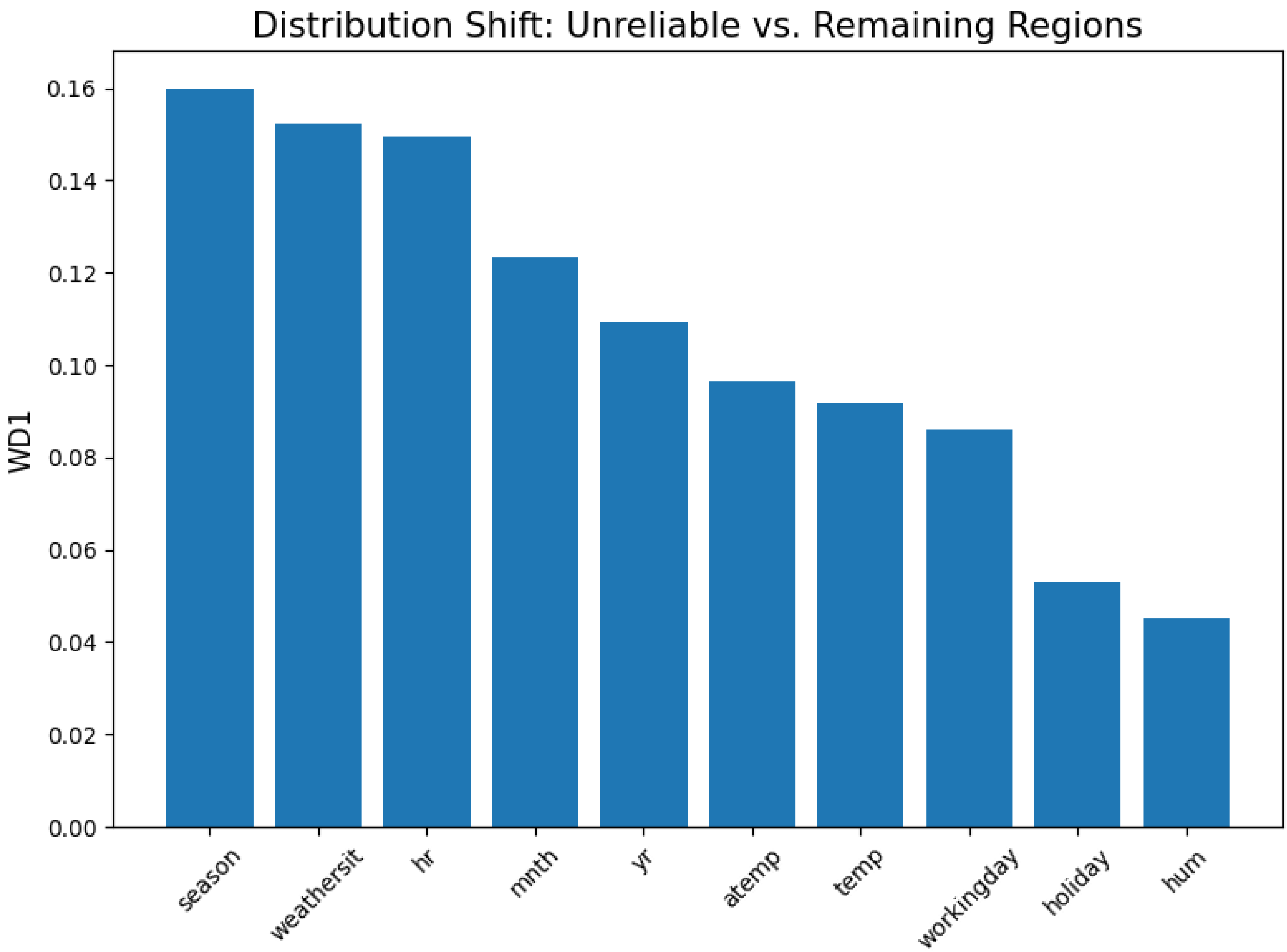


- #### Setting a threshold redefined the region that is considered unreliable changing the impact on the distribution. What was defined as unreliable or not in the previous graphs has changed creating higher PSI values.

Q7 (2.5 pts): Do the same analysis as in Q6, but use the Wasserstein distance measure instead of PSI. Are the results the same?

In [9]:

```
exp_bike_sharing.model_diagnose(model="XGB_Bike_Sharing", show="reliability_distance", threshold= 1.3, distance_metric='WD1')
```

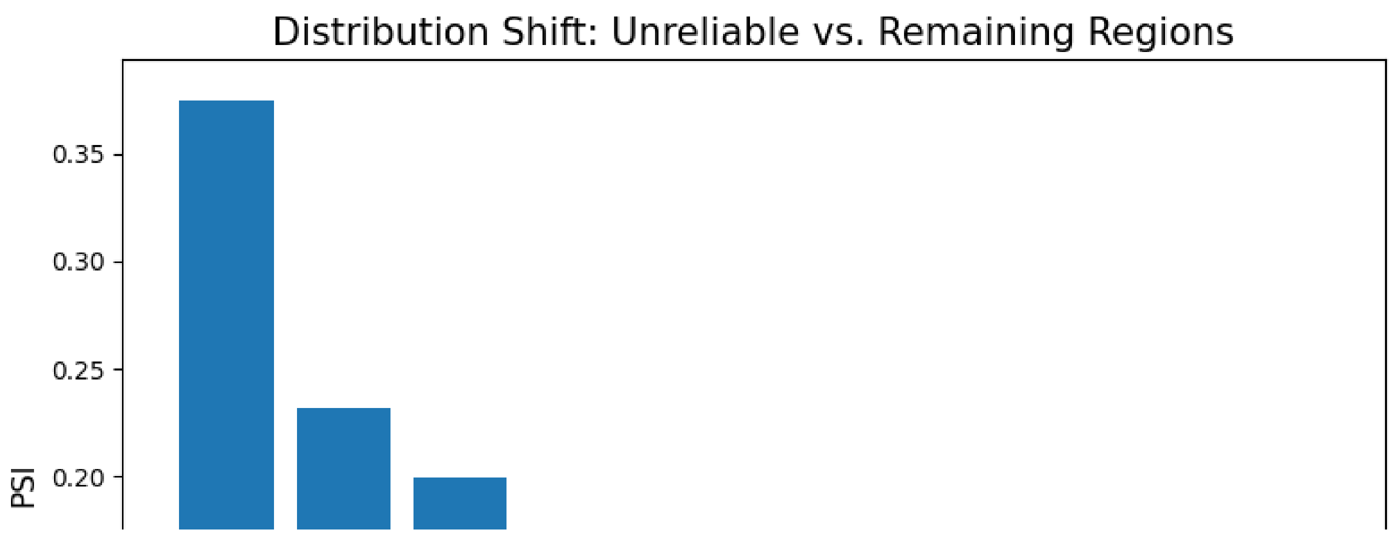


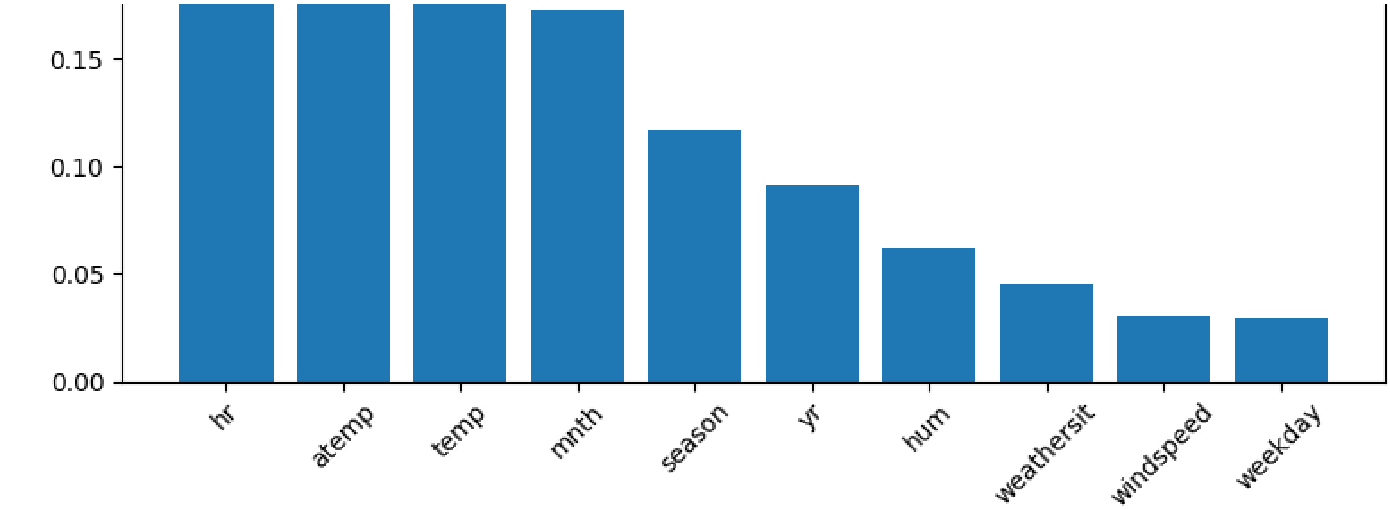
- #### No they aren't the same because both metrics capture different types of distribution shifts. PSI identifies features with significant changes in their distribution proportions while Wasserstein Distance highlights features that experience a broader shift in distribution shape.

Q8 (5 pts): Do the same analysis as in Q5, but for the DNN model. Are the results the same as Q5? Why or why not?

In [10]:

```
exp_bike_sharing.model_diagnose(model="DNN_Bike_Sharing", show="reliability_distance")
```





- #### They are not the same, DNN has more sensitivity to hr compared to the XGB model. This could be because of how DNN structure/complexity shifting the data to capture more time-dependent patterns compared to XGB tree splitting algorithmn which specialized in easier splitting terms. DNN also likely handled nonlinear relationships better than XGB.

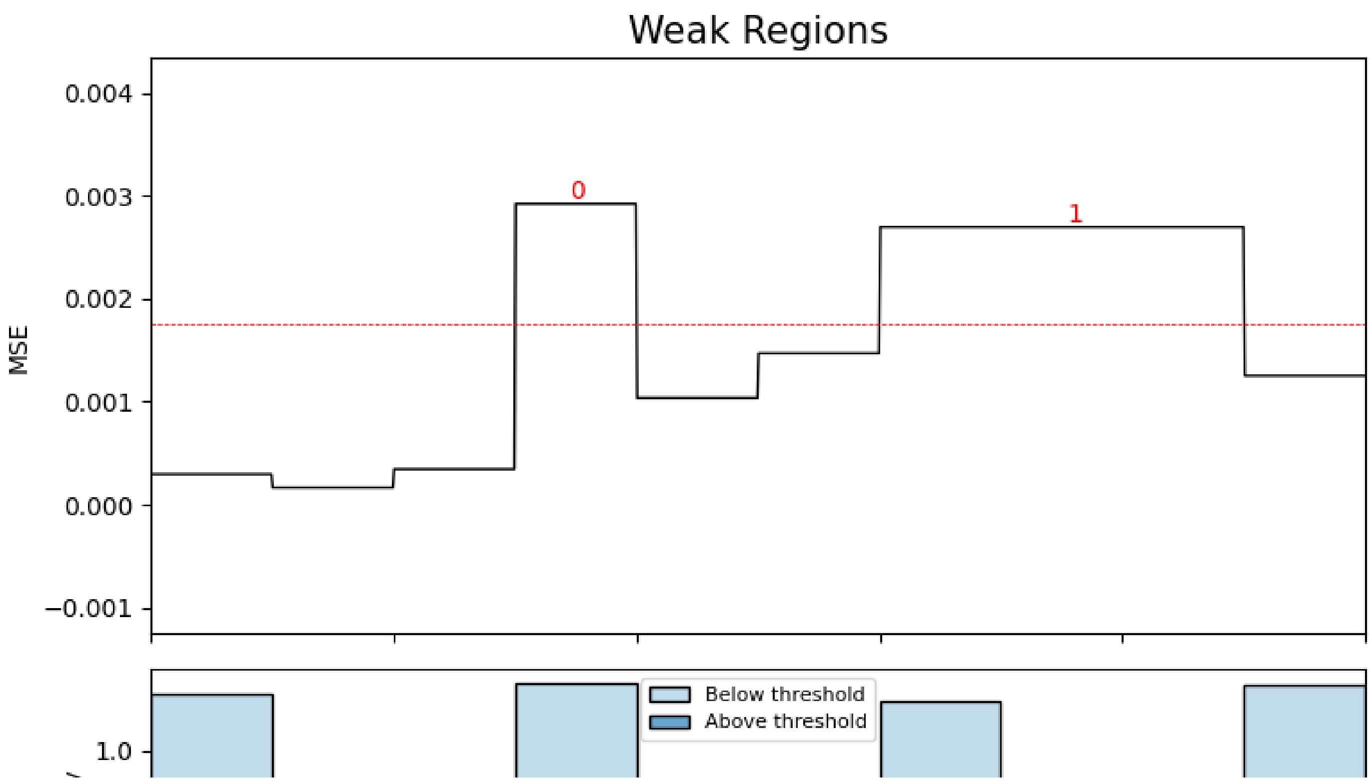
Q9 (5 pts): How could the analysis above help you if you were in charge of monitoring a model in production?

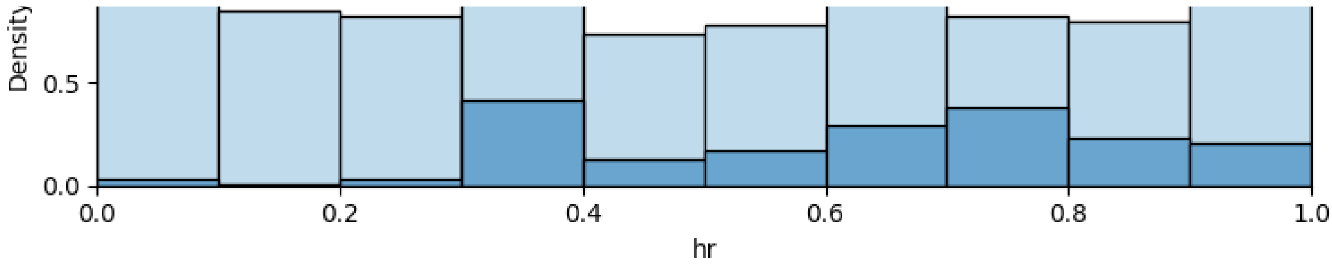
- #### These could help detect how severe our model is drifting via our data or feature distirbutions. They could be show us which features are causing the model drift and the overall stability of our model over a time period. Specifying which featurs and monitoring can ensure fairness and decrease biases.

Q10 (5 pts): For the `XGB_Bike_Sharing` model, write the PiML code to show the weak regions of the `hr` feature. Use MSE as the measurement metric and include the test data in the results. Use the `histogram` method for slicing.

In [11]:

```
exp_bike_sharing.model_diagnose(model="XGB_Bike_Sharing", show="weakspot", slice_method="histogram", slice_features=["hr"], metric="MSE", use_test=True)
```





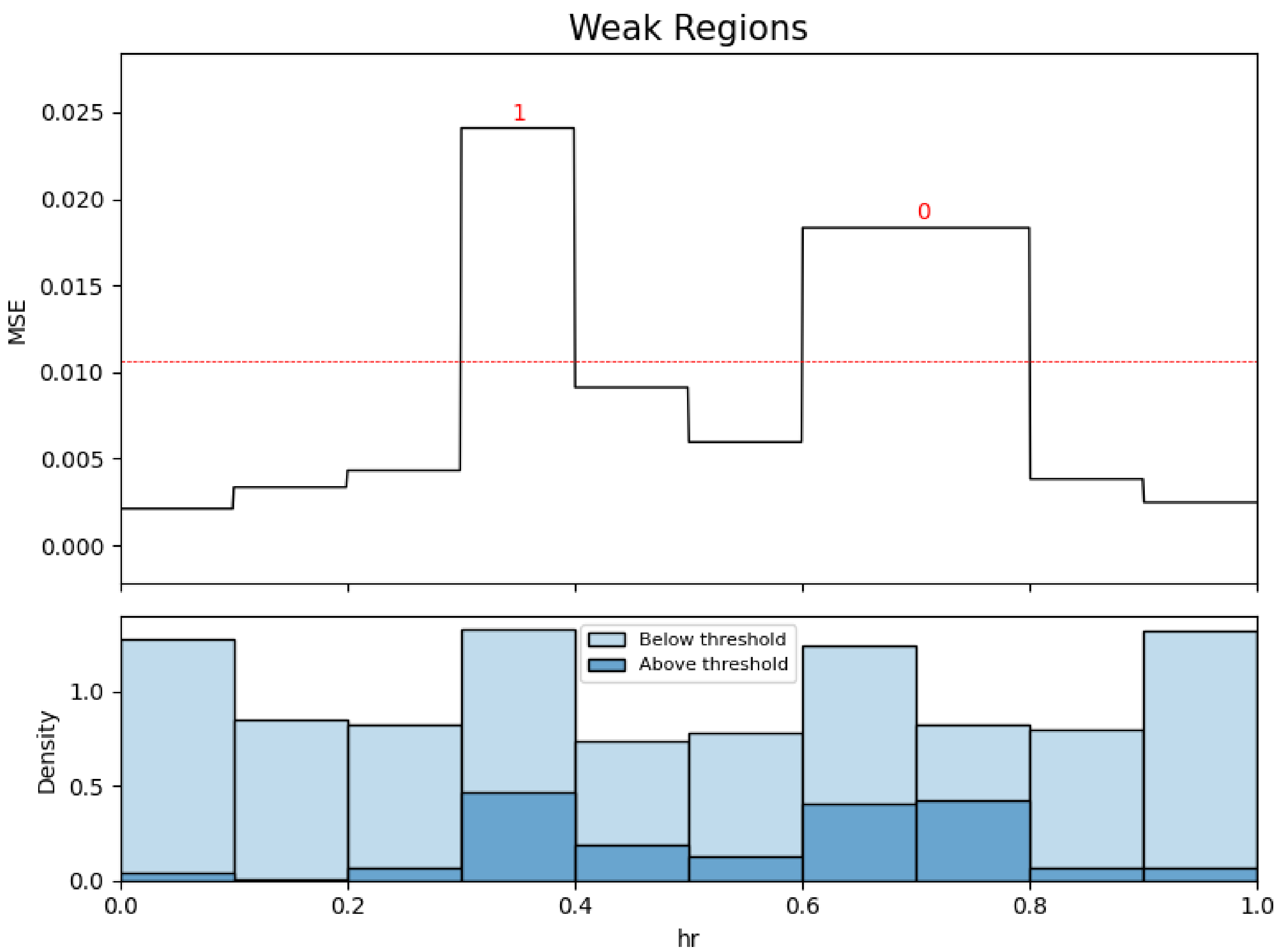
Q11 (5 pts): Is hour 3 or hour 4 part of a weak region in the analysis of Q10?

- #### Based on the plot, most of both hour 3 and 4 are above the threshold or having a high MSE meaning they are not in weak regions.

Q12 (5 pts): If you do the same analysis of Q10 for the DNN model, does it have the same weak regions for the `hr` feature?

In [12]:

```
exp_bike_sharing.model_diagnose(model="DNN_Bike_Sharing", show="weakspot", slice_method="histogram", slice_features=["hr"], metric="MSE", use_test=True)
```



- #### The DNN model does not have the exact same weak regions as the XGB model for the hr feature. DNN has more pronounced errors around certain hours compared to the XGB model, which might be capturing the temporal patterns differently.

Q13 (5 pts): Do these weak regions mean that the model should not be used? What might

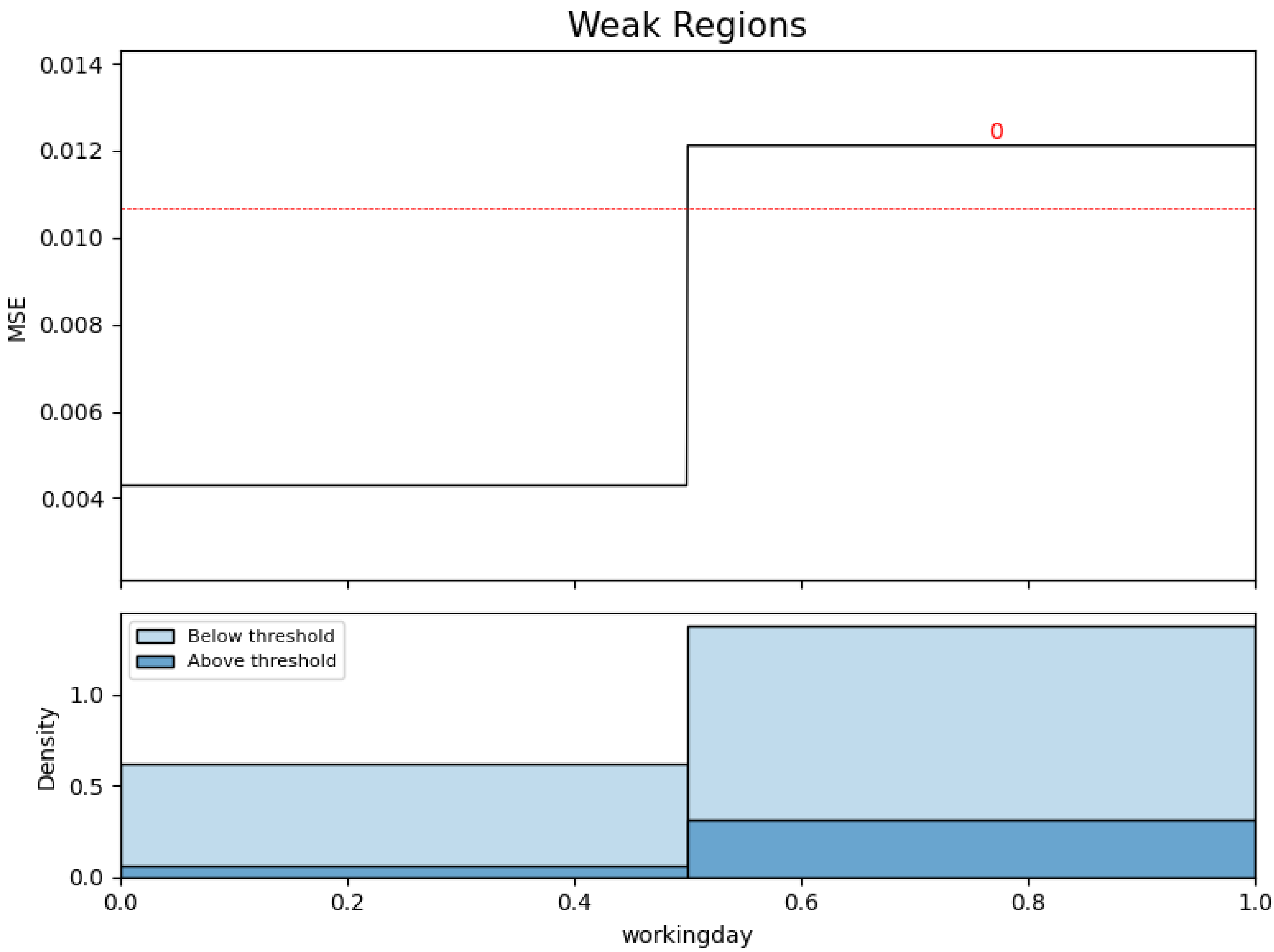
be done to improve the model in these regions?

- #### Just because the model has weak regions doesn't mean it shouldn't be used. Those spots defines specific areas are less reliable to model predictions during specific time periods and not generalizing well.
- #### Some things to improve it could be like feature engineering or categorization of some variables like time, adding additional data to those weak regions or using resampling strategies. The parameters also could be tuned for better generalization.

Q14 (10 pts): For the DNN model on the bike sharing data, analyze the weak spots for the `workingday` feature, based on testing data and MSE metric and historgram slicing. What is the number of test samples in the weak region? What is the difference from the test samples in the weak region to the MSE of the test data in the overall model?

In [17]:

```
exp_bike_sharing.model_diagnose(model="DNN_Bike_Sharing", show="weakspot", slice_method="
histogram", slice_features=["workingday"], metric="MSE", use_test=True, return_data=True
).data
```



Out[17]:

	[workingday	workingday)	#Test	#Train	test_MSE	train_MSE	Gap
0	0.5	1.0	2400	9465	0.012121	0.01191	0.000211

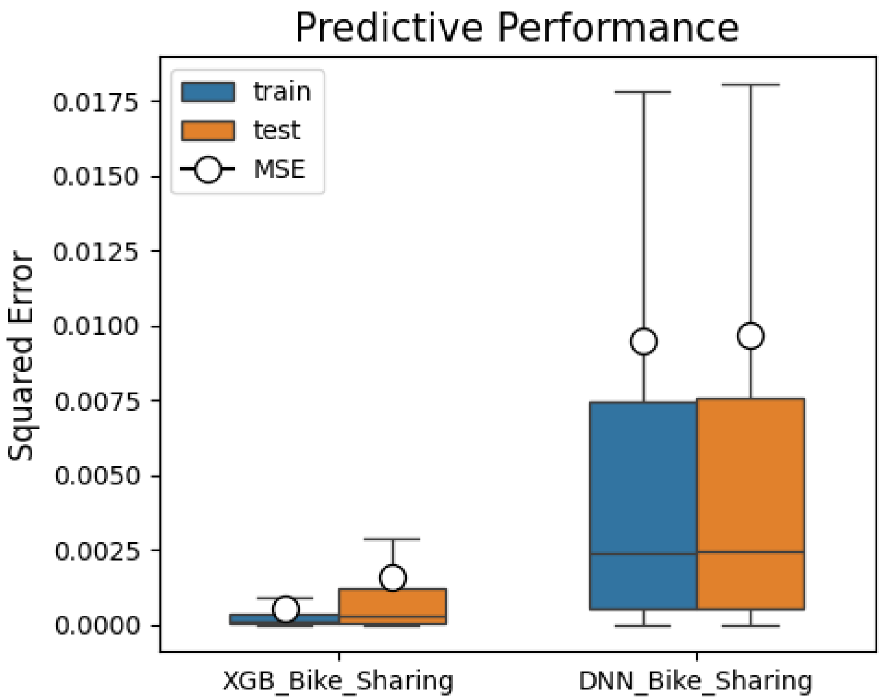
- #### Number of test samples in the weak region is 2400 (approximately 50% of the data). The weak region for the workingday feature has an MSE that is 0.002121 higher than the overall model test MSE. This indicates that the model struggles slightly more with predictions when workingday is within the interval from 0.5 to 1.0.

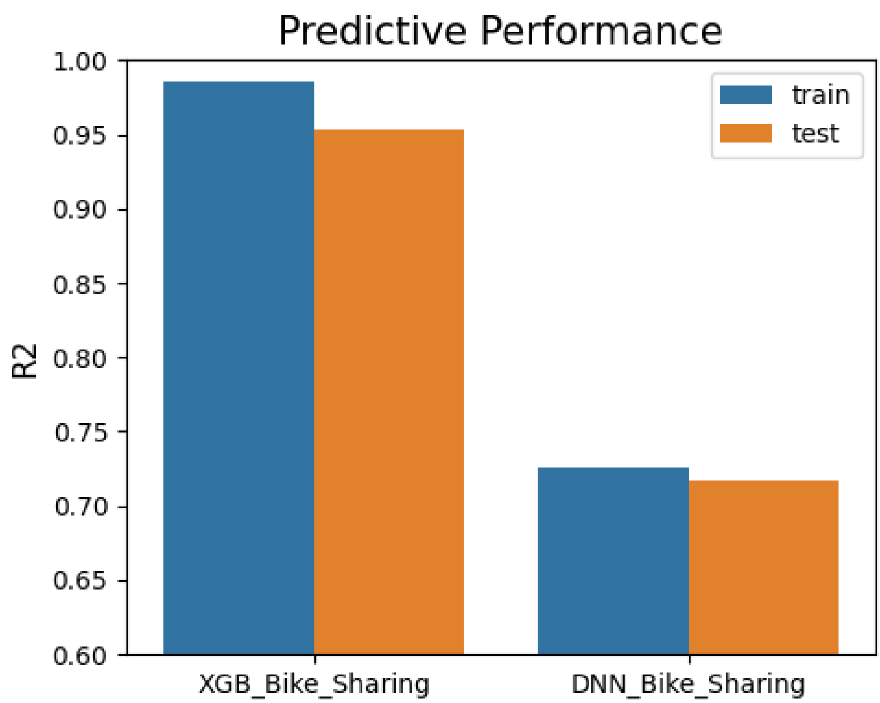
Q15 (10 pts): Compare the MSE and R2 for the two models used in biking sharing data set, for both training sample and testing. Plot the box plot for MSE and bar plot for R2. Which model performs better based on out-of-sample evaluation metrics?

In [14]:

```
exp_bike_sharing.model_compare(models=["XGB_Bike_Sharing", "DNN_Bike_Sharing"], show="acc
uracy_plot", metric="MSE", figsize=(5, 4))

exp_bike_sharing.model_compare(models=["XGB_Bike_Sharing", "DNN_Bike_Sharing"], show="acc
uracy_plot", metric="R2", figsize=(5, 4))
```





- **#### Based on the size of the box-plots and predictive performance on the bar graph, the XGB model is perfomring leaps ahead compared to the DNN. XGB has lower mean squared error and higher R^2, likely meaning that it is better at generalizing the data overall and making more accurate predictions on the unseen data.**

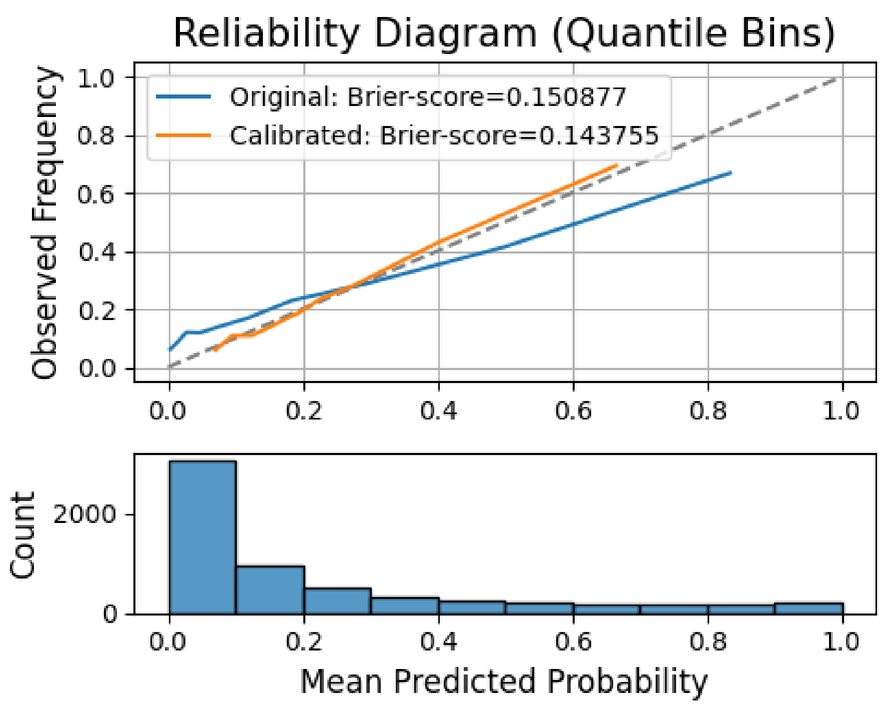
Taiwan (TW) credit dataset with XGB model and DNN model

Q16 (10 pts): Different from regression models, what may need to be calibrated in a classification model? Use PiML to plot the reliability diagrams (or calibration curves) for both of the Taiwan credit data set models. Then add a copy of that plot with 100 bins. Which model is more reliable before calibration for the Taiwan credit data set?

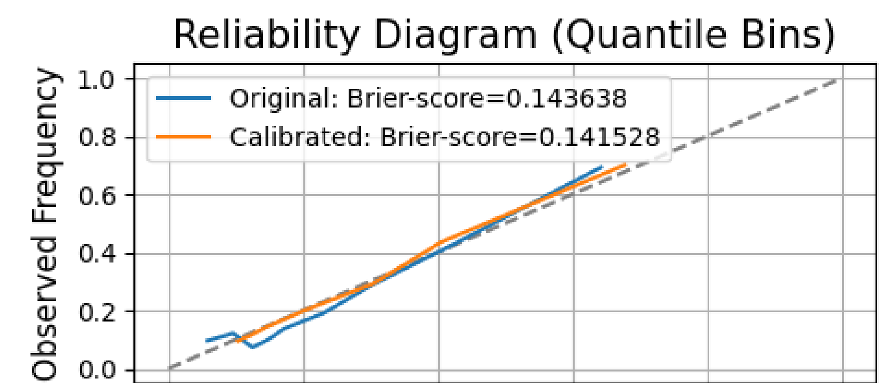
In [15]:

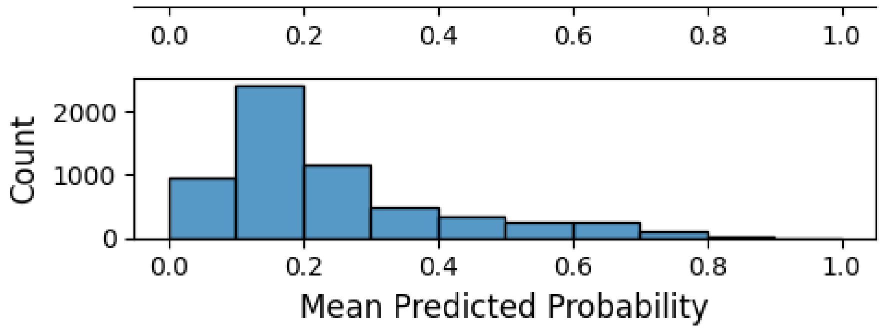
```
print("**************************************************************************")
print("Reliability Diagram XGB Taiwan Credit")
exp_tw_credit.model_diagnose(model="XGB_TW_Credit", show="reliability_perf", figsize=(5,
4))
print("**************************************************************************")
print("Reliability Diagram DNN Taiwan Credit")
exp_tw_credit.model_diagnose(model="DNN_TW_Credit", show="reliability_perf", figsize=(5,
4))
print("**************************************************************************")
print("Reliability Diagram XGB Taiwan Credit 100 Bins")
exp_tw_credit.model_diagnose(model="XGB_TW_Credit", show="reliability_perf", bins=100, fi
gsize=(5, 4))
print("**************************************************************************")
print("Reliability Diagram DNN Taiwan Credit 100 Bins")
exp_tw_credit.model_diagnose(model="DNN_TW_Credit", show="reliability_perf", bins=100, fi
gsize=(5, 4))
print("**************************************************************************")
```

```
**************************************************************************
Reliability Diagram XGB Taiwan Credit
```

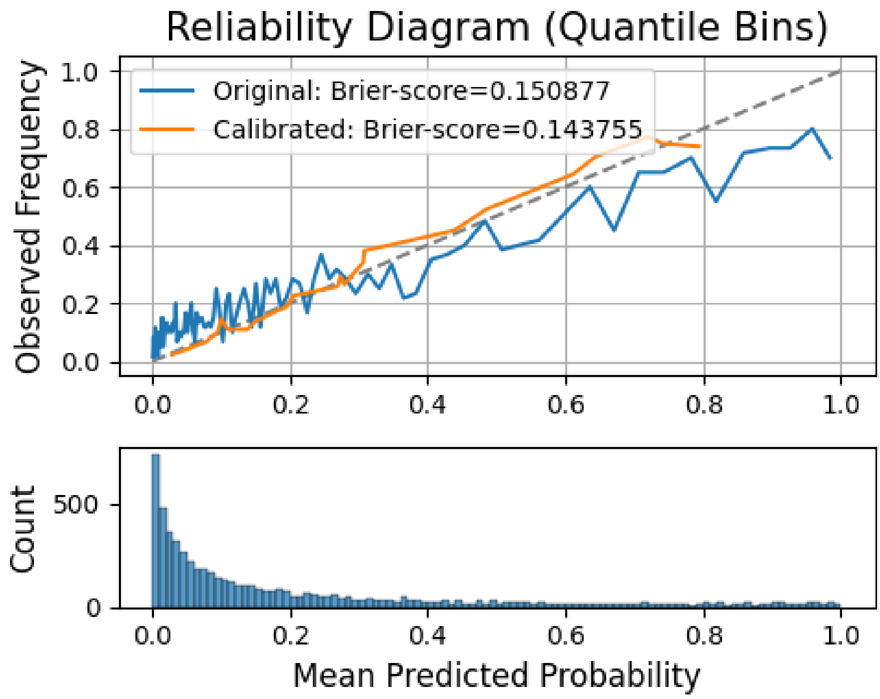



```
**************************************************************************
Reliability Diagram DNN Taiwan Credit
```

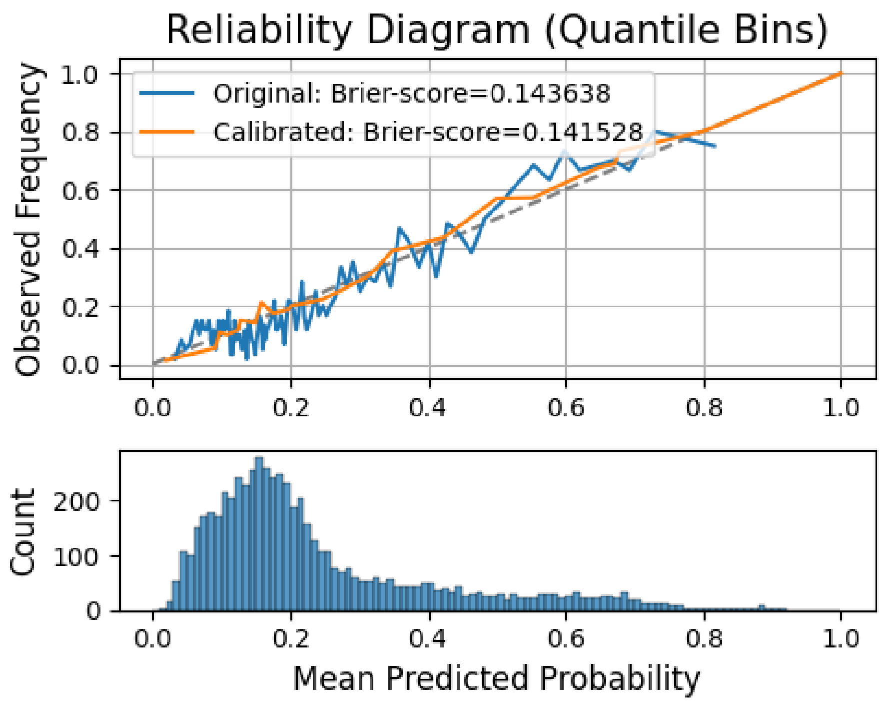





```
**********************************************************************
Reliability Diagram XGB Taiwan Credit 100 Bins
```




```
**********************************************************************
Reliability Diagram DNN Taiwan Credit 100 Bins
```




```
**********************************************************************
```

- **#### What needs to be calibrated are two main things, the predicted probabilities to ensure the likelihood of our data is accurate, and the decision threshold to deal with the imbalanced classes.**

- #### The model that is more accurate before the 100 bins is DNN. This is indicated by the calibration curve being closer to the ideal diagonal line and having a slightly better brier score than the XGB model, indicating the predicted probabilities better reflect the observed frequencies.

Q17 (5 pts): For the two models of TW credit dataset, which model performs better and why? Does any model potentially have overfitting problem and why?

- #### I think the DNN model performs better for the TW Credit set, mostly because the calibration line and brier scores were better, lower Brier scores indicate better calibrated probability estimates and the line was almost diagonal.
- #### The model that has the potential to overfit, this is seen in its calibration line being not well calibrated (either being to rigid or straying from the diagonal line), creating a situation where the model itself is not generalizing well. Tree-based algorithmns like XGB can overfit if the depth or the number of trees is too much, causing it to capture noise in the training data.

In [ ]:

```

```